Journal of General Microbiology (1987), **133**, 3403–3413. *Printed in Great Britain*

Localization of Plasmid Loci Necessary for the Entry of *Shigella flexneri* into HeLa Cells, and Characterization of One Locus Encoding Four Immunogenic Polypeptides

By BERNADETTE BAUDRY,[1] ANTHONY T. MAURELLI,[1]† PHILIPPE CLERC,[1] JERALD C. SADOFF[2] AND PHILIPPE J. SANSONETTI[1]*

[1] *Service des Entérobactéries, Unité INSERM 199, Institut Pasteur, 75724 Paris Cedex 15, France*

[2] *Department of Bacterial Diseases, Walter Reed Army Institute of Research, Washington, DC 20307-1500, USA*

(*Received 15 June 1987; revised 11 August 1987*)

We have previously cloned a 44 kb fragment from the virulence plasmid of *Shigella flexneri* serotype 5 strain M90T which is capable of restoring invasiveness to an avirulent, plasmidless mutant. This report presents a genetic and physical analysis of Tn*5* mutations in recombinant clone pHS4108. Tn*5* mutagenesis allowed identification of at least five regions implicated in the entry phenotype. These regions were located on a 20 kb portion of pHS4108. Expression of the insertion mutants was studied by immunoblots using the serum of a convalescent monkey infected by *S. flexneri* 2a, which recognized four plasmid-associated polypeptides. We propose that the four immunogenic polypeptides, a, b, c and d, are encoded by an operon.

INTRODUCTION

The pathogenic potential of *Shigella flexneri* is correlated with the ability of this organism to enter, multiply within and kill human colonic epithelial cells (La Brec *et al.*, 1964). Genetic studies have established that in addition to several chromosomal loci (Formal & Hornick, 1978), a large plasmid of 220 kb is necessary for expression of this invasive phenotype (Sansonetti *et al.*, 1982). Although complete expression of virulence, as measured by a positive Sereny test (Sereny, 1955), requires chromosomal loci, the transfer of pWR100, the 220 kb plasmid of *S. flexneri* serotype 5 strain M90T, to an *Escherichia coli* K12 strain demonstrated that the plasmid by itself is sufficient to promote invasion of HeLa cells (Sansonetti *et al.*, 1983).

In an attempt to characterize the genetic sequences as well as the molecular products necessary for invasion of eukaryotic cells by *S. flexneri*, we have cloned pWR100 invasion sequences into the cosmid vector pJB8 (Ish-Horowicz & Burke, 1981). Several recombinant plasmids containing a common sequence of 37 kb carried the genetic information necessary to restore the invasive ability of an avirulent, plasmidless mutant of *S. flexneri* into HeLa cells (Maurelli *et al.*, 1985). However, a complete invasive phenotype was not expressed by the recombinant clones in that they failed to produce a positive Sereny test. Further experiments demonstrated that they were impaired in their ability to replicate within eukaryotic cells, thus indicating that other regions of the plasmid were also important (Sansonetti *et al.*, 1986). Such an hypothesis is also supported by recent data of Sasakawa *et al.* (1986*a*, *b*), who showed that several distinct regions are implicated in the positive Sereny phenotype and that insertion in loci as far as 100 kb apart gave a negative Sereny test. However, study of our recombinant plasmids has enabled us to focus on the genetics and molecular basis of the entry process itself.

† Present address: Department of Microbiology, Uniformed Services University of the Health Sciences, Bethesda, MD 20814-4799, USA.

Table 1. *Bacterial strains, plasmids and phage*

Strain or phage	Relevant characteristics	Plasmid content	Invasion of HeLa cells	Source or reference
Shigella flexneri 5				
M90T	Wild-type, *nad*	220 kb virulence plasmid pWR100	+	Sansonetti *et al.* (1983)
M90TminI	Minicell producer mutant of M90T	220 kb virulence plasmid pWR100	+	Hale *et al.* (1983)
M90TminII-34	Spontaneous plasmidless derivative of M90TminI	No virulence plasmid	−	Hale *et al.* (1983)
Shigella flexneri 2a				
BS169	Mal$^+$ λ^s *galU*::Tn*10*	No virulence plasmid	−	Maurelli *et al.* (1985)
BS169/pHS4108	Ap^r	45 kb insert from pWR100 cloned into pJB8	+	Maurelli *et al.* (1985)
BS123	λpapa lysogen of BS169	No virulence plasmid	−	This work
Escherichia coli				
HB101	F$^-$ *ara-14 leu proA2 lacY1 glnV44* λ^- *galK2 recA13 rpsL20 xyl-5 mtl thi hsdS20*	None	−	Boyer & Roulland-Dussoix (1969)
Phage				
λ467	*λb221 rex*::Tn*5 cI857* $O_{am}29$ $P_{am}80$			Berg *et al.* (1975)

Preliminary studies of protein expression by the recombinant plasmids have shown that among the polypeptides expressed by this long sequence of DNA, four were consistently recognized by sera of monkeys immunized against *S. flexneri* (Maurelli *et al.*, 1985). These polypeptides, which have been called a (78 kDa), b (62 kDa), c (43 kDa) and d (38 kDa) (Hale *et al.*, 1985), represent the major proteinaceous antigens which elicit a humoral response during infection of monkeys and human beings by *Shigella* species (Oaks *et al.*, 1986). However, immunogenicity of these four polypeptides did not imply that they were directly involved in the entry process. In addition, one expected to find other entry-associated polypeptides that may not be immunogenic.

This report describes the study of one of these recombinant plasmids, pHS4108. It contains a 44 kb insert which we have mutagenized using transposon Tn*5*. Eight different insertions of the transposon appeared to inhibit or alter the entry process. They mapped within scattered loci on a 20 kb portion of the cloned insert. We then focused on the region encoding the four immunogenic polypeptides. The data suggested that they were encoded by an operon.

METHODS

Bacterial strains, plasmids and phage. These are listed in Table 1.

Media and culture conditions. Bacteria were routinely grown in tryptic soy broth (Institut Pasteur Production) or L Broth (Lennox, 1955). For infection with phage λ, strains were grown in L broth (no glucose) plus 10 mM-$MgSO_4$. Antibiotics were added to broth culture or brain heart infusion (BHI) agar (Difco) in the following concentrations: ampicillin, 100 μg ml^{-1}; kanamycin, 40 μg ml^{-1}.

Isolation, characterization and cloning of plasmid DNA. For transformation and screening of plasmid DNA, 1·5 ml overnight broth cultures were extracted by the technique of Ish-Horowicz & Burke (1981). DNA for restriction analysis was further purified by phenol extraction. Restriction endonuclease digestion of plasmid DNA was performed according to the recommendations of the manufacturers. Electrophoresis was carried out in 0·7% (w/v) agarose gels with E buffer (40 mM-Tris, 2 mM-disodium EDTA, pH 7·9) as running buffer. Cloning experiments were performed with dephosphorylated vector DNA as described by Maniatis *et al.* (1982).

Insertion mutagenesis of pHS4108. A λ467 lysate was used to infect, at a multiplicity of infection of 5, 1 ml of an exponential-phase culture of *S. flexneri* BS213/pHS4108 grown in the presence of ampicillin. After 20 min adsorption at 37 °C, the infected culture was shaken at 37 °C for 45 min, then diluted with 9 ml L broth containing kanamycin and incubated overnight at 37 °C. Plasmid DNA extracted from 1·5 ml of the resulting overnight culture was used to transform *E. coli* HB101 and Ap^{r}Km^r transformants were selected. An average of 10 transformants per mutagenesis experiment was chosen for analysis. The plasmids carried by these clones were studied by restriction analysis and transformed into *S. flexneri* BS169 for infection of HeLa cells.

Infection of HeLa cells. Non-confluent monolayers of HeLa cells in 35 mm plastic tissue culture dishes (Becton Dickinson Labware) were inoculated with bacteria, centrifuged for 10 min at 2200 *g* and incubated for 2 h at 37 °C. Monolayers were then washed, Giemsa-stained and examined by light microscopy for invasion of cells (Hale & Formal, 1981). A strain was considered invasion$^+$ (Inv$^+$) when at least 30% of the cells were infected. Strains referred to as Inv$^\pm$ showed invasion of less than 5% of the cells. Strains were considered Inv$^-$ when several infection assays showed no infected cells.

Contact haemolysis assay. Assay for haemolytic activity was performed as previously described (Clerc *et al.*, 1986; Sansonetti *et al.*, 1986).

Preparation of anti-polypeptide b monoclonal antibodies. The band corresponding to polypeptide b was cut out from a preparative SDS gel. After electroelution, this polypeptide was injected to BALB/c mice. The protocol for immunization, fusion with the mouse myeloma cell line X63-Ag8.653 and cloning by limiting dilution has already been described (Sadoff *et al.*, 1985). The antigen used to screen hybridomas which produced antibodies against polypeptide b was a water extract of *S. flexneri* M90T (Oaks *et al.*, 1986). The hybridoma cell line producing the 1.3.358.59.2 monoclonal antibody was selected. Ascites fluid of mice was used in immunoblot studies.

Western blot hybridization. Whole bacterial extracts were run on SDS-polyacrylamide gels and blotted onto nitrocellulose filters as described by Burnette (1981). Diluted (1/250) convalescent serum from a monkey which had been infected with *S. flexneri* 2a was used to detect expression of four major polypeptides encoded by the virulence plasmid of M90T (Hale *et al.*, 1985) and by recombinant plasmid pHS4108 (Maurelli *et al.*, 1985).

Analysis of proteins expressed in minicells. Purification of minicells from 14 h BHI cultures was accomplished by differential centrifugation and three sucrose density gradient separations (Gemski & Griffin, 1980). Purified minicells were labelled for 2 h with [^{35}S]methionine [50 μCi ml^{-1} (1·85 MBq ml^{-1}), >800 Ci mmol^{-1}, Amersham] as previously described (Gill *et al.*, 1979). After washing, minicells were solubilized at 37 °C for 30 min in a 8 M-urea solution containing 2% (w/v) octyl-β-D-glucopyranoside (Sigma), 2% (w/v) pH 3·5–9·5 Ampholine (LKB) and 5% (v/v) 2-mercaptoethanol (Hale *et al.*, 1985). Remaining cellular debris was discarded after centrifugation and approximately 10^6 c.p.m. of TCA-precipitable material was analysed by two-dimensional gel electrophoresis (O'Farrell *et al.*, 1977). Polypeptides were resolved by non-equilibrium pH gradient electrophoresis in flat-bed polyacrylamide gels consisting of 5% (w/v) acrylamide, 2% (v/v) Triton X-100, 2% (w/v) pH 3·5–9·5 premixed Ampholine (Hale *et al.*, 1985) and 8 M-urea. Gels were submitted to a 1 h pre-run at 30 mA constant current before addition of the samples and migration for 2000 V h. A linear pH gradient of approximately 4 to 9·5 was established. Individual tracks were then excised and frozen at −70 °C. For separation in the second dimension SDS-PAGE, sample tracks were equilibrated for 1 h in two changes of SDS sample buffer (O'Farrell, 1975) before insertion on a vertical slab gel consisting of a 5% (w/v) polyacrylamide stacking gel and a 12·5% (w/v) running gel (Laemmli, 1970). Fixation and fluorography of dried gels was performed as previously described (Hale *et al.*, 1983).

RESULTS

Mutagenesis of pHS4108 and characterization of insertion mutants

To identify regions on pHS4108 essential for invasion of HeLa cells, insertions of Tn*5* in the plasmid were isolated using a λ::Tn*5* donor. In order to avoid λTer-mediated recombination between λ::Tn*5* and the cos site of pHS4108 (De Bruijn & Lupski, 1984), the recipient *S. flexneri* strain, BS169, was first lysogenized with λpapa. This lysogenic strain, referred to as BS213, was then transformed with pHS4108 and mutagenized as described in Methods. A total of 76 Ap^r Km^r insertion mutants was studied. Tn*5* insertions, pHS4108::Tn*5*-1 to pHS4108::Tn*5*-76, were first roughly localized using the restriction enzyme *Eco*RI, which has no cleavage site within Tn*5* (Jorgensen *et al.*, 1979), then precisely mapped by *Sal*I, *Bam*HI and *Xho*I restriction analysis.

The assay for invasiveness of HeLa cells showed that eight sites of insertion scattered along a 20 kb internal segment of the insert corresponded to an altered entry phenotype. As shown in Fig. 1, three insertions each separated from the next by at least 3 kb on the central part of the insert (insertions 53, 37 and 29), and two close insertions (31 and 1) mapping between kb 32 and

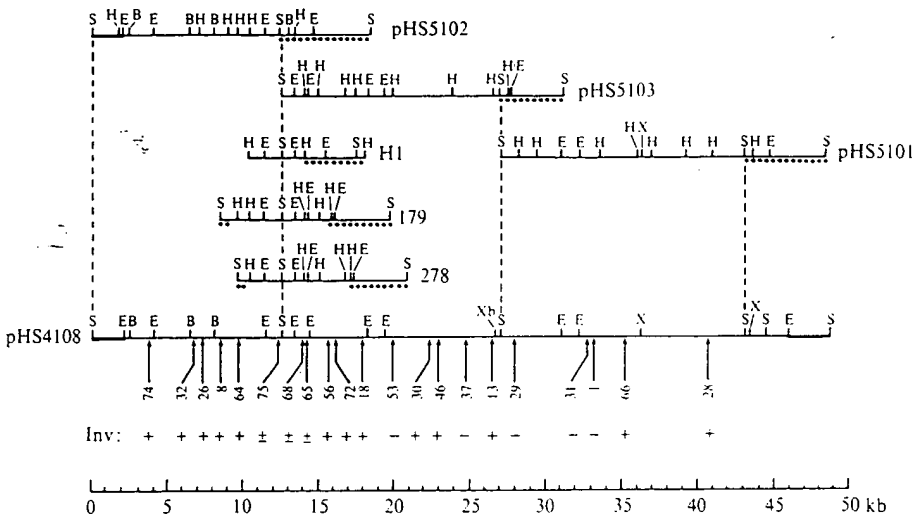

Fig. 1. Map of the recombinant plasmid pHS4108 with localization of Tn*5* insertions. For each Tn*5* insertion presented, the corresponding entry phenotype (Inv$^+$, $\geqslant 30\%$ of cells infected; Inv$^\pm$, $< 5\%$ of cells infected; Inv$^-$, non-invasive) is indicated. In the upper part are represented the three *Sal*I fragments cloned from pHS4108, the *Hin*dIII subclone H1 and *Sau*3A subclones 278 and 179, which expressed immunogenic polypeptides. Thick lines represent vector pJB8. Other cloning vectors are underlined with dots. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I; X, *Xho*I; Xb, *Xba*I.

33 on the right-hand end of the sequence, corresponded to a total loss of penetration capacity. We previously reported that strains harbouring pHS4108 did not express contact haemolytic activity (Sansonetti *et al.*, 1986). This was certainly a consequence of the instability of the cosmid and of the poor selection provided by ampicillin. In the case of Tn*5* mutants, selection by kanamycin enabled us to measure this ability. We observed that all the non-invasive mutants had lost contact haemolytic activity while all insertion mutants which were still invasive expressed haemolytic activity at approximately 50% of the level of wild-type strain M90T. The other three independent insertions altering the entry process mapped within 2 kb (between 12 and 14 kb on the scale): insertions 65 and 68 clustered in less than 0·5 kb, and insertion 75 located 1·5 kb farther to the left. These three insertion mutants showed a tenfold decrease in the rate of HeLa cell infection, as compared with BS169/pHS4108, and the contact haemolytic activity of mutants 65 and 68 was also decreased (contact haemolytic activity was not determined for mutant 75). All these results suggested that at least five different loci, separated by non-essential sequences, were involved in the entry process.

Expression of immunogenic polypeptides by Tn5 mutants and by subclones

pWR100 and pHS4108 encode four plasmid-associated polypeptides recognized by serum of monkeys which have been orally infected by *S. flexneri* (see Introduction). In order to determine if some insertion mutants failed to express one or more of these polypeptides, immunoblots of whole-cell extracts were performed.

Five independent mutants were found to have modified expression of some or all of the four polypeptides. These mutants harboured plasmids pHS4108 : : Tn*5*-8, -64, -65, -68 and -75. The Tn*5* insertions carried by these plasmids defined a region of about 7 kb (see Figs 1 and 4). Clones BS169/pHS4108 : : Tn*5*-8 and BS169/pHS4108 : : Tn*5*-64 (Fig. 2, lane 8) did not express polypeptide a. Insertion mutants 65 and 68 showed a dramatic decrease in expression of polypeptides a, b, and d, and a slight diminution of polypeptide c (Fig. 2, lanes 5 and 6).

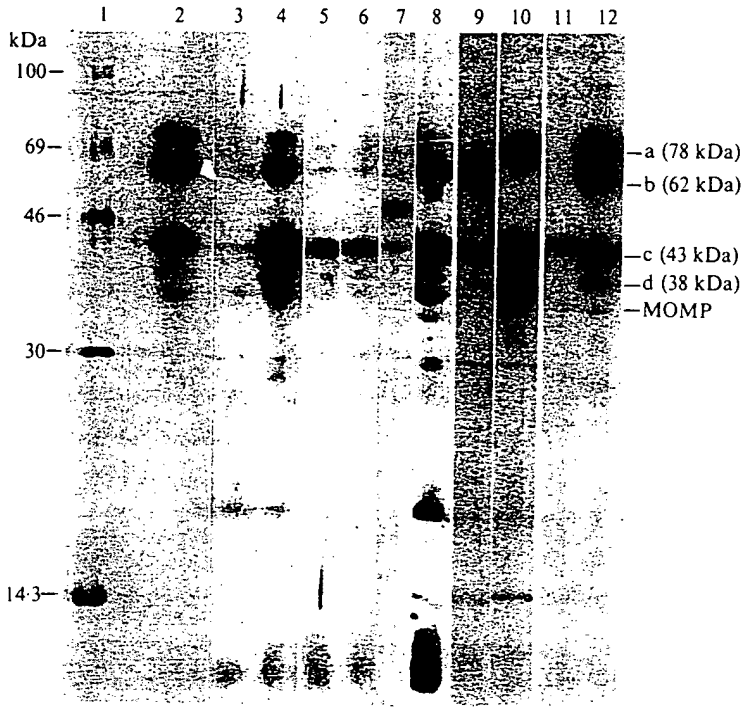

Fig. 2. Immunoblot of polypeptides expressed by the insertion mutants. Total protein extracts were obtained from strain BS169 carrying the mutagenized plasmids. Lanes: 1, molecular mass standards; 2, M90T; 3, BS169; 4, mutant 32; 5, mutant 65; 6, mutant 68; 7, mutant 75; 8, mutant 64; 9, *Sau*3A subclone 278; 10, BS169/pHS5102; 11, BS169/pHS5103; 12, M90T. MOMP is a major outer-membrane protein.

Insertion mutant 75 showed weak expression of polypeptides a and d, a slight decrease in expression of c and expressed a new polypeptide of approximately 48 kDa (Fig. 2, lane 7). All the other insertion mutants, whether invasive or not, showed expression of the polypeptides quantitatively similar to that of the wild-type strain M90T (see insertion mutant 32, Fig. 2, lane 4).

Strains containing plasmids pHS5101, pHS5102, pHS5103 (respectively 16, 12·5 and 14·5 kb *Sal*I fragments of pHS4108 cloned into pBR322 or pBR325: Fig. 1) were also tested in immunoblots for expression of the polypeptides. BS169/pHS5101 did not express any of the four polypeptides. BS169/pHS5102 showed expression of polypeptides a, c and d similar to the control, M90T (Fig. 2, lane 10). BS169/pHS5103 expressed only one immunogenic polypeptide, of approximately 43 kDa (Fig. 2, lane 11).

In order to determine the origin of this 43 kDa polypeptide, and to confirm that polypeptide b was not expressed by pHS5102, protein extracts of BS169/pHS5102 and BS169/pHS5103 were immunoblotted with a monoclonal antibody which recognizes an epitope of polypeptide b. As expected, the 43 kDa polypeptide encoded by pHS5103 was recognized by this anti-b monoclonal antibody (Fig. 3, lane 3), thus showing that it was a truncated product of the gene encoding polypeptide b. The results obtained with clone BS169/pHS5102 confirmed that no polypeptides were recognized by the anti-b monoclonal antibody.

Simultaneously, fragments of 5 to 10 kb from a *Sau*3A partial digestion of plasmid pHS4108 had been cloned into the *Bam*HI restriction site of pBR322. Relying on the immunoblot data obtained with the insertion mutants, subclones containing the 1·2 kb *Eco*RI fragment (between

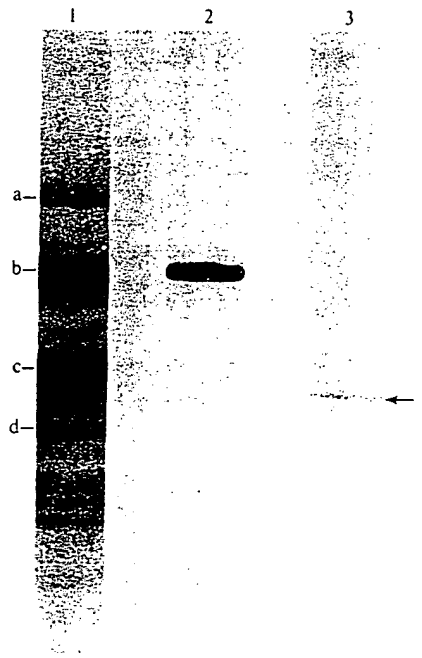

Fig. 3. Immunoblot with the anti-b monoclonal antibody. Lanes: 1, immunoblot of whole-cell extract of M90T with a convalescent monkey serum showing polypeptides a, b, c and d; 2, immunoblot of the same M90T extract with mouse ascites fluid containing the anti-b monoclonal antibody 1.3.358.59.2; 3, BS169/pHS5103 lysate reacted with the anti-b monoclonal antibody (the arrow indicates the truncated polypeptide b encoded by pHS5103).

13 and 14·5 kb on the scale in Fig. 1) were selected by colony hybridization using the *Eco*RI 1·2 kb::Tn*5* fragment of insertion mutant 68 as a probe. Protein extracts from the selected subclones were tested in immunoblots with the serum. Two subclones, 278 and 179, were found to express polypeptides b, c and d (Fig. 2, lane 9). Maps of these two subclones are shown in Fig. 1.

In another attempt to subclone some of the genes encoding these polypeptides, the 4·5 kb *Hin*dIII fragment located between kb 10 and 15 (see Fig. 1) was cloned into the *Hin*dIII restriction site of pACYC184. Subclones in each of the two possible orientations of the insert within the vector were selected. After transformation of the plasmids into M90TminII-34, whole-cell extracts were tested in immunoblots for expression of the antigenic polypeptides. Expression of polypeptides b and c was detected. All the subclones expressing b and c were found to have their insert in the same orientation (see H1, Fig. 1).

According to these data, a map of the genes encoding polypeptides a, b, c and d was constructed (Fig. 4).

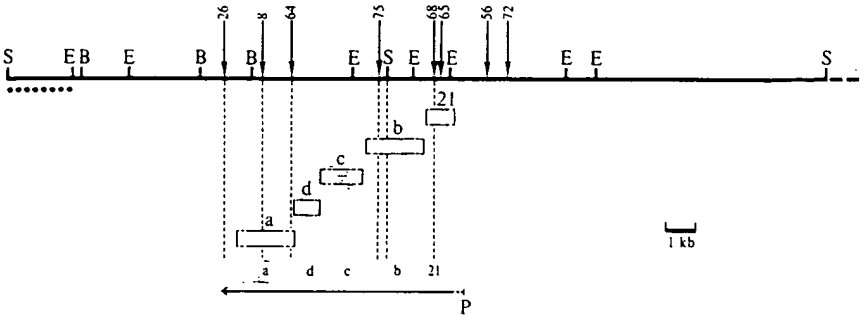

Fig. 4. Part of the physical map of pHS4108 with the location of the genes encoding immunogenic polypeptides a, b, c and d and the 21 kDa polypeptide of pI 9. The vertical arrows correspond to the Tn*5* insertions. The ends of the open boxes corresponding to the genes are represented by broken lines because the precise limits of the genes are not known. The horizontal arrow represents the direction of transcription from the predicted promoter P.

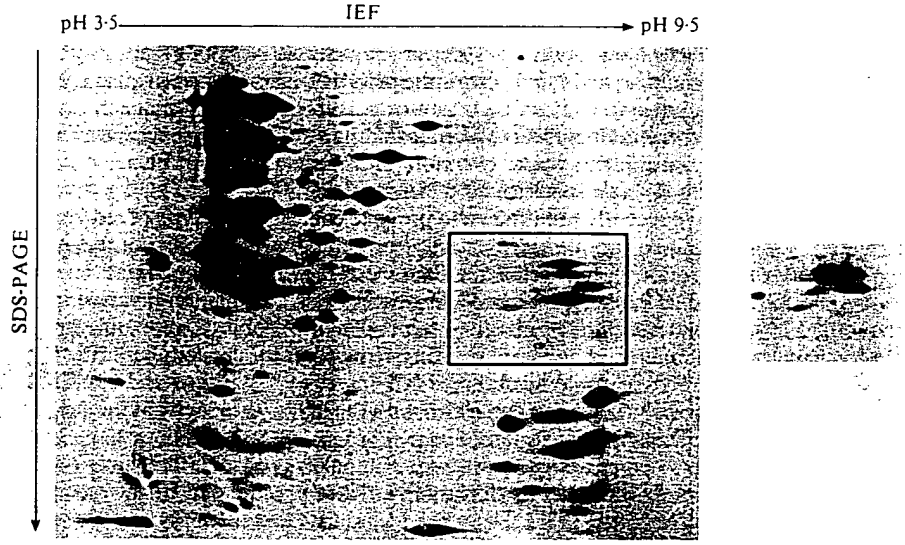

Fig. 5. Two-dimensional electrophoresis gel of ^{35}S-labelled protein from M90TminII-34/pHS4108 : : Tn*5*-50, an invasive insertion mutant in which the transposon is inserted in the vector. A detail of the two-dimensional profile of M90TminII-34/pHS4108 : : Tn*5*-65 is shown; the asterisk emphasizes the absence of the 21 kDa polypeptide of pI9 as compared with the framed area in the main photograph. Radiolabelled products were detected on X-ray film after 48 h exposure.

Two-dimensional gel electrophoresis of proteins

Although they showed a decrease in expression of the immunogenic polypeptides a to d, insertion mutants 65 and 68 showed no disappearance or modification of these polypeptides. In an attempt to identify additional polypeptides, two-dimensional gel electrophoresis of ^{35}S-labelled proteins from minicells of insertion mutant 65 was performed. The two-dimensional profile of insertion mutant 65 was compared with that of M90TminII-34/pHS4108 : : Tn*5*-50, which is an invasive insertion mutant carrying the transposon on the vector. One polypeptide of 21 kDa with a pI of 9 could no longer be detected on the two-dimensional profile of insertion mutant 65 (Fig. 5). The presence of this polypeptide on the two-dimensional profile of the

control M90TminI and its absence on the two-dimensional profile of M90TminII-34/pJB8 : : Tn*5* confirmed that the 21 kDa polypeptide of pI 9 was indeed encoded by a gene of the insert of pHS4108.

DISCUSSION

Invasion of epithelial cells of the large intestine is an essential feature of the pathogenicity of *S. flexneri* and requires expression of genes located on the 220 kb virulence plasmid pWR100 (Sansonetti *et al.*, 1982). Cosmid cloning of pWR100 DNA demonstrated that a sequence of 37 kb, contained in pHS4108, encoded all the information necessary for entry into epithelial cells (Maurelli *et al.*, 1985). The next step was to localize the entry-associated genes within this sequence on recombinant plasmid pHS4108.

Tn*5* mutagenesis of pHS4108 enabled us to identify five non-contiguous regions in which insertions abolished or altered the entry phenotype. These insertions were localized within a 20 kb internal region of pHS4108. In addition, none of the subclones containing either one of the three largest *Sal*I fragments, or one of the *Eco*RI fragments of 14·0 kb, 11·5 kb or 7·6 kb, restored the invasive ability (data not shown). These results confirmed that a plasmid sequence of at least 20 kb was necessary for entry of *S. flexneri* into epithelial cells. Watanabe & Nakamura (1986) have characterized a plasmid sequence of 20 kb which is necessary for invasion in *Shigella sonnei*. Comparison of restriction maps shows a striking similarity between the *S. sonnei* sequence and the right-hand part of the pHS4108 insert, between 20 and 39 kb on the scale in Fig. 1. In addition, in one of their non-invasive Tn*1* insertion mutants, cloned on plasmid pHW511, the transposon appears to be located in a site very close to that of the Tn*5* mutants 1 and 31 described in this work. Therefore, it appears likely that invasion genes are highly conserved among *Shigella* species.

Invasive shigellae have recently been shown to provoke contact haemolysis of red blood cells (Clerc *et al.*, 1986). This plasmid-mediated activity correlates with the ability of the strains to enter HeLa cells, and to multiply intracellularly (Sansonetti *et al.*, 1986). The contact haemolytic activity of pHS4108 : : Tn*5* mutants matched this observation. All the non-invasive insertion mutants had lost this activity while mutants retaining invasive ability were still capable of contact haemolysis. Moreover, in mutants showing reduced entry capacity, a corresponding decrease of haemolytic activity could be observed (i.e. insertion mutants 65 and 68). Thus, not only was the contact haemolytic activity correlated with the invasive capacity of the strains, but its level reflected the efficiency of invasion.

Hale *et al.* (1985) have demonstrated that serum of monkeys immunized against *S. flexneri* recognized four polypeptides encoded by the large virulence plasmid and by the recombinant plasmid pHS4108 (Maurelli *et al.*, 1985). These results indicate that these four polypeptides are major antigens. This was confirmed by results obtained using sera of patients recovering from shigellosis (Oaks *et al.*, 1986). The role of these polypeptides in the virulence of *Shigella* is unknown. However, our results with insertion mutants 8 and 64 clearly showed that polypeptide a was not implicated in the entry process, at least in the HeLa cell model. On the other hand, the fact that insertion mutants showing reduced expression of the polypeptides also expressed an altered entry phenotype indicates that an least one of these polypeptides plays an important role in the invasion process. In these mutants (i.e. 65 and 68), expression of polypeptide b in particular was dramatically decreased. It is possible that the lack of polypeptide b may be responsible for the decreased penetration efficiency. The fact that insertion mutant 75, which produced a truncated polypeptide b, expressed the same reduced invasion phenotype as insertion mutants 65 and 68 would favour this possibility.

Analysis of immunoblots performed on the insertion mutants and on the subclones allowed the construction of a genetic map of the sequence encoding the four polypeptides (Fig. 4). Because simultaneously altered expression of the four polypeptides can be obtained with one Tn*5* insertion, we suggest that the genes encoding the immunogenic polypeptides are clustered in an operon.

Analysis of two-dimensional gel electrophoresis profiles allowed identification of the product

of the gene inactivated in insertion mutant 65. The polypeptide had a molecular mass of 21 kDa and a pI of 9, and probably corresponds to the g polypeptide identified as a virulence-associated plasmid product by Hale *et al.* (1985). As discussed below, there is a strong presumption that this polypeptide is also encoded by a gene of the operon.

Based on all the data, a model for the genetic organization of the operon is proposed. The operon would contain five genes encoding polypeptides a, b, c, and d and the 21 kDa/pI 9 polypeptide. The direction of transcription would be from right to left (i.e. 21 kDa/pI 9, b, c, d and a) according to the orientation of the map (Fig. 4). The promoter should be located within the space defined by the sites of insertion in mutants 65 and 56, that is, between 14·5 and 16 kb on the scale drawn on Fig. 1. This is based on the facts that insertion 65 altered expression of the five polypeptides and insertion 56 had no effect either on expression of the polypeptides or on the invasive ability. The fact that the cloned *Hin*dIII 4·5 kb fragment expressed polypeptides b and c only in one orientation favoured this hypothesis: it showed that the promoter of the operon was not on the 4·5 kb *Hin*dIII fragment, and that in these clones expression of the polypeptides was under the control of a promoter located on the vector. The weak but detectable expression of the genes when located downstream from the site of insertion could be explained by residual transcription through Tn*5*. The occurrence of this phenomenon when Tn*5* is used to mutagenize an operon carried by a multicopy plasmid has already been described (De Bruijn & Lupski, 1984). Immunoblots also showed that expression of polypeptide c was more efficient than that of the other polypeptides when Tn*5* insertions were located upstream, within the genes encoding b or the 21 kDa/pI 9 polypeptide. The better expression of polypeptide c could be explained either by a more efficient ribosome-binding site in front of the gene encoding polypeptide c, or by the presence of a secondary promoter at the beginning of this gene. Alternatively, the 21 kDa/pI 9 polypeptide could be a positive regulatory protein which would activate the operon encoding polypeptides a, b, c and d. The highly basic property of this polypeptide would be consistent with a DNA-binding protein.

Buysse *et al.* (1987) have recently described the cloning, in the expression vector λgt11, of fragments from the *S. flexneri* virulence plasmid encoding some of the immunogenic polypeptides. The clones expressed either polypeptides b, or c, or d, or b and c. Based on the maps of these clones, a map of the region encoding polypeptides b, c and d has been constructed by these authors. This map perfectly matches the data described in this work, in the position of the restriction sites and in the order of the genes encoding polypeptides b, c and d on the sequence. However, although no definitive data are presented, the authors favour the hypothesis of a regulon model, that is independent genes regulated by a common activator.

We are currently subcloning the region encoding polypeptides a, b, c and d, in order to study the genetic organization of the genes for these four polypeptides more precisely and determine whether the gene encoding the 21 kDa/pI 9 polypeptide is indeed included in the operon. We are also in the process of subcloning all the Tn*5* insertions corresponding to a non-invasive phenotype, to identify other entry-associated polypeptides.

The available data indicate that the invasive ability of *S. flexneri* is a very complex phenomenon which involves many genes and a large array of polypeptides. The invasion strategy developed by *Shigella* species is certainly very different from that of *Yersinia pseudotuberculosis*, in which only one gene appears necessary for promoting entry of bacteria into cells (Isberg & Falkow, 1985). In the case of *S. flexneri*, we have shown that at least five distinct genes or clusters of genes are required for the entry step. Whether all these gene products are directly involved in the interaction with the cells, or whether a pool of polypeptides is necessary for transformation and/or correct positioning of a unique product is not yet known.

We wish to thank Thomas L. Hale for kindly providing sera of monkeys immunized against *S. flexneri* and for helpful discussions throughout this work. We also wish to thank Pascale Courcoux and Hélène d'Hauteville for excellent technical assistance.

This work was supported by a grant from the Diarrhoeal Disease Control Programme of the World Health Organisation. B.B. was under a fellowship of the Ministère de la Recherche et de l'Enseignement Supérieur. A.T.M. was supported by a postdoctoral fellowship from the European Molecular Biology Organization and by a grant from the Association pour le Développement de l'Institut Pasteur.

REFERENCES

BERG, D. E., DAVIES, J., ALLET, B. & ROCHAIX, J. D. (1975). Transposition of R factor genes to bacteriophage lamba. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 3628–3632.

BOYER, H. W. & ROULLAND-DUSSOIX, D. (1969). A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *Journal of Molecular Biology* **41**, 459–472.

BURNETTE, W. N. (1981). 'Western blotting': electrophoretic transfer of proteins from sodium dodecyl sulfate polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radio-iodinated protein A. *Analytical Biochemistry* **112**, 195–203.

BUYSSE, J. M., STOVER, C. K., OAKS, E. V., VENKATESAN, M. & KOPECKO, D. J. (1987). Molecular cloning of invasion plasmid antigen (*ipa*) genes from *Shigella flexneri*: analysis of the *ipa* gene products and genetic mapping. *Journal of Bacteriology* **169**, 2561–2569.

CLERC, P., BAUDRY, B. & SANSONETTI, P. J. (1986). Plasmid-mediated contact hemolytic activity in *Shigella* species: correlation with penetration into HeLa cells. *Annales de l'Institut Pasteur/Microbiologie* **134A**, 267–278.

DE BRUIJN, F. J. & LUPSKI, J. R. (1984). The use of transposon Tn*5* mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids – a review. *Gene* **27**, 131–149.

FORMAL, S. B. & HORNICK, R. B. (1978). Invasive *Escherichia coli*. *Journal of Infectious Diseases* **137**, 641–644.

GEMSKI, P., JR & GRIFFIN D. E. (1980). Isolation and characterization of minicell-producing mutants of *Shigella* spp. *Infection and Immunity* **30**, 104–111.

GILL, R. E., HEFFRON, F. & FALKOW, S. (1979). Identification of the protein encoded by the transposable element Tn*3* which is required for its transposition. *Nature, London* **282**, 797–801.

HALE, T. L. & FORMAL, S. B. (1981). Protein synthesis in HeLa and Henle407 cells infected with *Shigella dysenteriae* 1, *Shigella flexneri* 2a, or *Salmonella typhimurium* W118. *Infection and Immunity* **32**, 137–144.

HALE, T. L., SANSONETTI, P. J., SCHAD, P. A., AUSTIN, S. & FORMAL, S. B. (1983). Characterization of virulence plasmids and virulence-associated outer membrane proteins in *Shigella flexneri*, *Shigella sonnei* and *Escherichia coli*. *Infection and Immunity* **40**, 340–350.

HALE, T. L., OAKS, E. V. & FORMAL, S. B. (1985). Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infection and Immunity* **50**, 620–629.

ISBERG, R. R. & FALKOW, S. (1985). A single genetic locus encoded by *Yersinia pseudotuberculosis* permits invasion of cultured animal cells by *Escherichia coli* K-12. *Nature, London* **317**, 262–264.

ISH-HOROWICZ, D. & BURKE, J. D. (1981). Rapid and efficient cosmid vector cloning. *Nucleic Acids Research* **9**, 2989–2998.

JORGENSEN, R. A., ROTHSTEIN, S. J. & REZNIKOFF, W. S. (1979). A restriction enzyme cleavage map of Tn*5* and localization of a region encoding neomycin resistance. *Molecular and General Genetics* **177**, 65–72.

LA BREC, E. H., SCHNEIDER, H., MAGNANI, T. J. & FORMAL, S. B. (1964). Epithelial cell penetration as an essential step in the pathogenesis of bacillary dysentery. *Journal of Bacteriology* **88**, 1503–1518.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

LENNOX, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**, 190–206.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

MAURELLI, A. T., BAUDRY, B., D'HAUTEVILLE, H., HALE, T. L. & SANSONETTI, P. J. (1985). Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infection and Immunity* **49**, 164–171.

OAKS, E. V., HALE, T. L. & FORMAL, S. B. (1986). Serum immune response to *Shigella* protein antigens in Rhesus monkeys and humans infected with *Shigella* spp. *Infection and Immunity* **53**, 57–63.

O'FARRELL, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry* **250**, 4007–4021.

O'FARRELL, P. Z., GOODMAN, H. M. & O'FARRELL, P. H. (1977). High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* **12**, 1133–1142.

SADOFF, J. C., WRIGHT, D. C., FUTROVSKY, S., SIDBERRY, H., COLLINS, H. & KAUFMANN, B. (1985). Characterization of mouse monoclonal antibodies directed against. *Pseudomonas aeruginosa* lipopolysaccharides. In *Antibiotics and Chemotherapy*, vol. **36**, *Pseudomonas aeruginosa: New Therapeutic Approaches from the Basic Research*, pp. 134–167. Edited by D. P. Speert & R. E. W. Hancock. Basel: Karger.

SANSONETTI, P. J., KOPECKO, D. J. & FORMAL, S. B. (1982). Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infection and Immunity* **35**, 852–860.

SANSONETTI, P. J., HALE, T. L., DAMMIN, G. J., KAPFER, C., COLLINS, H. H., JR & FORMAL, S. B. (1983). Alteration of the pathogenicity of *Escherichia coli* K12 after transfer of plasmid and chromosomal genes from *Shigella flexneri*. *Infection and Immunity* **39**, 1392–1402.

SANSONETTI, P. J., RYTER, A., CLERC, P., MAURELLI, A. T. & MOUNIER, J. (1986). Multiplication of *Shigella flexneri* within HeLa cells: lysis of the phagocytic vacuole and plasmid-mediated contact hemolysis. *Infection and Immunity* **51**, 461–469.

SASAKAWA, C., KAMATA, K., SAKAI, T., MURAYAMA, S. Y., MAKINO, S. & YOSHIKAWA, M. (1986*a*). Molecular alteration of the 140 megadalton plasmid associated with the loss of virulence and Congo red binding activity in *Shigella flexneri*. *Infection and Immunity* **51**, 470–475.

SASAKAWA, C., MAKINO, S., KAMATA, K. & YOHIKAWA, M. (1986*b*). Isolation, characterization, and mapping of Tn*5* insertions into the 140 megadalton invasion plasmid defective in the mouse Sereny test in *Shigella flexneri* 2a. *Infection and Immunity* **54**, 32–36.

SERENY, B. (1955). Experimental *Shigella* conjunctivitis. *Acta microbiologica Academiae scientiarum hungaricae* **2**, 293–296.

WATANABE, H. & NAKAMURA, A. (1986). Identification of *Shigella sonnei* form I plasmid genes necessary for cell invasion and their conservation among *Shigella* species and enteroinvasive *Escherichia coli*. *Infection and Immunity* **53**, 352–358.